

1/24

Figure 1:

Query= INSP087
(1357 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
1,039,285 sequences; 328,747,273 total letters

Searching.....done

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| ref XP_090334.2 (XM_090334) similar to ovostatin precursor - ch... | 1747 | 0.0 |
| ref XP_132895.1 (XM_132895) similar to ovomacroglobulin, ovosta... | 1707 | 0.0 |
| emb CAA55385.1 (X78801) ovomacroglobulin, ovostatin [Gallus gal... | 1170 | 0.0 |
| sp P20740 OVOS_CHICK OVOSTATIN PRECURSOR (OVOMACROGLOBULIN) >gi ... | 1170 | 0.0 |
| ref NP_000005.1 (NM_000014) alpha 2 macroglobulin precursor [Ho... | 995 | 0.0 |
| ref NP_036620.1 (NM_012488) alpha-2-macroglobulin [Rattus norve... | 990 | 0.0 |
| prf 1009174A macroglobulin alpha2 [Homo sapiens] | 986 | 0.0 |
| ref NP_002855.1 (NM_002864) pregnancy-zone protein; Pregnancy z... | 965 | 0.0 |
| pir JC5143 alpha-macroglobulin precursor - guinea pig >gi 13040... | 963 | 0.0 |
| ref NP_665722.1 (NM_145779) pregnancy-zone protein [Rattus norv... | 942 | 0.0 |

Note: The top two matches are XP_090334.2 and XP_132895.1 which were predicted by computational automated analysis using a gene prediction method.

2/24

Figure 2:

```

>emb|CAA55385.1| (X78801) ovomacroglobulin, ovostatin [Gallus gallus]
Length = 1454

Score = 1170 bits (3027), Expect = 0.0
Identities = 615/1369 (44%), Positives = 896/1369 (64%), Gaps = 31/1369 (2%)

Query: 1      VPQARSDPLAFITFSAKGATLNLEERRSVAIRSRENVVFVQTDKPTYKPGQKVHILTLFL 60
          +P   S  LAFI+F+AKG T +L+ERRSV I + E+ VFVQTDKP YKPGQ V      + L
Sbjct: 84     IPPVTSVSLAFISFTAKGTTFDLKERRSVMINMESFVVFVQTDKPIYKPGQSVMFVVAL 143

Query: 61     -FLFQ-----YPVITLQDPQNNRIFQRQNVTSFRNITQLSFQLISEPMFGDYWIVVKRNS 114
          F F+      YP+I +QDPQNNRIFQ QNVTS  NI Q+ F L  EP+ G+Y I+V + S
Sbjct: 144     DFNFKPVQEMYPLIAVQDPQNNRIFQWQNVTSIEINIVQIEFPLTEEPILGNYKIIIVTKKS 203

Query: 115    RETVTHQFAVKRYVLPKFETVNAAPQTVTISDDEFQVDVCAKYNFGQPVQGETQIRVCRE 174
          E  +H F V+ YVLPKF+VTV AP ++T+ D E  V +CA, Y +GQPV+G+ Q+ VCR+
Sbjct: 204     GERTSHSFLVEEYVLPKFDTVTAPGSLTVMDSELTVKICAVYTYGQPVEGKVQLSVCRD 263

Query: 175    YFSSSNCEKNEENEICEQFIAQLE-NGCVSQIVNTKVFQLYRSGLFMTFHVAVIVTESGT 233
          + S  C+K+  +C+ F  L+ +GC+S I+++KVF+L R G      V  IVTE  V
Sbjct: 264     FDSYGRCKKSP--VCQSFITKDLTDGCLSHILSSKVFELNRIGYKRNLDVKAIIVTEKEQV 321

Query: 234    MQISEKTSVFITQLLGTVNENMDTFYRRGISYFGTLKFSDPNNVPMVNKLQLLELNDEF 293
          ++  S+ ITQ++ ++ FEN+D  YRRGI YFG +K  D +N P+ NK++QL +N++
Sbjct: 322     CNLTATQSIISITQVMSSLQFENVDHHYRRGIPYFGQIKLVDKDNSPISNKVIQLFVNNKN 381

Query: 294    IGNYTTDENGEAQFSIDTSIDFPEFNLKATYVRPESCYLPSWLTPQYLDAHFLVSRFYS 353
          N+TTD NG A FSIOTS IFDPE +LKA Y  + C+  W+ P Y DA  V R YS
Sbjct: 382     THNFTTDINGIAPFSIDTSKIFDPELSLKALYKTSDQCHSEGWIEPSYPDASLSVQLRLYS 441

Query: 354    RTNSFLKIVPEPKQLECNQQKVVTVHYSLNSEAYEDDSNVKFFYLMMVKGAILLSGQKEI 413
          T+SF++I P  K + C Q++++TV+Y LN+E YE  + V F+Y+ M KG I+L+G+ ++
Sbjct: 442     WTSSFVRIEPLWKDMSCGQKRMITVYYILNTEGYEHINIVNFYYVGMAKGKIVLTGEIKV 501

Query: 414    RNKA-WNGNFSFPISISADLAPAAVLFVYTLHPSGEIVADSVRFQVDKCFKHKVNIFSN 472
          +A  NG F  P+ ++  +APA  L  VY LHP+ E+VADSVRF ++KCFK+KV ++FS
Sbjct: 502     NIQADQNGTFMIPLVVNEKMAPALRLLVYMLHPAKELVADSVRSIEKCFKNVKQLFSE 561

Query: 473    EQGLPGSNASLCLQAAPVLFCALRAVDRNVLLKSEQQLSAESVYNMVPSEPYGYFYHG 532
          +Q  L  SN  SL ++AA  FCA+RAVD++LLLKSE +LSAE++YN+ P  +  GY ++G
Sbjct: 562     KQMLTTSNVSLVIEAAAANSFCAVRAVDKSMLLLKSETELSAETIYNLHPIQDLQGYIFNG 621

```

Query: 533 LNLDGGKEDPCIPQRDMFYNGLYYPVSNYGDGDIYNIVRNMGGLKVFTNLHYRKPEVCVM 592

LNL+D +DPC+ D+F+ GLYY P++ D+Y +R+MG+K FTN R+P VC

Sbjct: 622 LNLEDDPQDPCVSSDDIFHKGLYYRPLTSGLGPDVYQFLRDMGMKFTNSKIRQPTVCTR 681

Query: 593 ERRLPLPKPLYLETENYGPMRSVPSRIACRGENADYVEQAI IQTVRTNF PETWMWDLVSV 652

E P P Y + + + + + I++T+R FPETW+WD++ +

Sbjct: 682 ETVRP---PSYFLNAGF-TASTHHVKLSAEVAREERGKRHILETIREFFPETWIWIDIILI 737

Query: 653 DSSGSANLSFLIPDTITQWEASGFCVNGDVGFGISSTTTLEVSQPFFIEIASPFSVVQNE 712

+S+G A++S+ IPDTIT+W+AS FCV GFG+S TL QPFF+++ P+S++ E

Sbjct: 738 NSTGKASVSYTIPDTITEWKASAFCVEELAGFGMSVPATLTAQFPFFVDLTLPYSIIHGE 797

Query: 713 QFDLIVNVFSYRNTCVEISVQVEESQNYEANIHTLKINGSEVIQAGGRKTNVWTIIPKKL 772

F + NVF+Y N C++I+V + ES +Y+A + + + +G + A RK+ VW I PK

Sbjct: 798 DFLVRANVFNLYNHCIKINVLLLESLDYQAKLISPEDDG--CVCAKIRKSYVWNIFPKGT 855

Query: 773 GKVNITVVAESKQSSACPNEGMEQQKLNWKDTVVQFLVEPEGIEKERTQSFLICTEGAK 832

G V ++ AE+ AC E + ++++++DT +++ LVEPEGI +E TQ+FLIC +

Sbjct: 856 GDVLFSITAETNDDEACEEEALRNIRIDYRDTQIRALLVEPEGIRREETQNFLICMKDDV 915

Query: 833 ASKQGVLDLNDVVEGSARGFFTVVGDILGLALQNL-VVLQMPYGSGEQNAALLASDTYV 891

S+ +DLP +VVEGS R F+VVGDI+G A+QN+ +LQMP+G+GEQN L A + YV

Sbjct: 916 ISQDVAIDLPTNVVEGSPRPSFSVVGDIMGTAIQNVHQLLQMPFGNGEQNMVLFAPNIYV 975

Query: 892 LDYLKSTEQLTEEVQSKAFFLLSNGYQRQLSFKNSDGSYSVFWQQSQKGS--ICALTFKT 949

LDYL T QL+E+V+SK L +GYQ+QLS+K+ DGSYS F + ++G+ + A +K+

Sbjct: 976 LDYLDKTRQLSEDVKSCTIGYLVSGYQKQLSYKHPDGSYSTFGIRDKEGNTWLTAFVYKS 1035

Query: 950 LERMKKVVFIDENVQKQTLIWLSSQQKTSGCFKNDGQLFNHAWEGLDEEDISLTAYVVM 1009

+++++ID+NVQ QTЛИWL+++QKT GCF++ G L N+A +GG E ++SL+AY+

Sbjct: 1036 FAEASRFIYIIDDNVQAQTLIWLATKQKTDGCFQSTGILVNNAMKGGVENELSLSAYITIA 1095

Query: 1010 FFEAGLNFTFPALRNALFCLEAALDSGVNTNGYNHAILAYAFALAGKEKQVESLLQTLQDS 1069

EAG + + +RNA +CLE A + +T+ Y A++AYAF LAGK + ES L+ L +S

Sbjct: 1096 LLEAGHSMSHTVIRNAFYCLETASEKNITDIYTQALVAYAFCLAGKAEICESFLRELQKS 1155

Query: 1070 APKLNNVIYWERERKPKTEEFPSFIPWAPSAQTEKSCYVLLAVI---SRKIPDLTYASKI 1126

A +++ YWE+ ++ E+ + S E + YVLLA++ +R DLT AS I

Sbjct: 1156 AKEVDGSKYWEQNQRSAPEK-SHLLDHVQSTDVEITSYVLLALLYKPNRSQEDLTKASAI 1214

Query: 1127 VQWLAQRMNHSGGFSSNQETAVCLLAITRYITQGLFSKDQNTVTFSEGSSE-IFQVNGH 1185

VQW+ ++ NS+GGF+S Q+T V L A+ Y S QN + +S+ + E +F VN

Sbjct: 1215 VQWIIIRQQNSYGGFASMQDTVVALQALAAYGAATYNSVTQNVKINSKNTFEKVFTVNNE 1274

Query: 1186 NRLLVQRSEVTQAPGEYTV DVEGHGCTFIQATLKYNVLLPKKASGFSL SLEIVKNYSST- 1244

NRLL+Q++ + Q PG+Y++ V G GC IQ L+YN+ LP+ A GFSL S++ N S

Sbjct: 1275 NRLLLQQTPLPQVPGKYSLTVNGTGCVLIQTALRYNIHLPEGAFGFSL SVQ-TSNASCPR 1333

Query: 1245 ----AFDLTVTLKYTGIRNKSSMVIDVKMLSGFTPTMSSIEELENKGQVMKTEVKNDHV 1300

FD+ + YTG R+ S+MV+IDVKMLSGF P SS+++L + VM+ E K +HV

Sbjct: 1334 DQPGKFDIVLISSYTGKRSSSNMVIIDVKMLSGFVVKSSL DQLIDDHTVMQVEYKKNHV 1393

Query: 1301 LFYLENGF-GRADSFPFSVEQSNLVFNIQPAPAMVYDYYEKEEYALAFY 1348

L YL N R FSVEQ +V + +PAP +YDYYE EEYA+A Y

Sbjct: 1394 LLYLGNILQKRRKEVTF SVEQDFVVTHPKPAPVQIYDYYETEEYAVA EY 1442

Figure 3:

Query= INSP088_pep
(894 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
1,039,285 sequences; 328,747,273 total letters

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|-----------------|------------|
| dbj BAC04793.1 (AK096448) unnamed protein product [Homo sapiens] | 974 | 0.0 |
| ref NP_000005.1 (NM_000014) alpha 2 macroglobulin precursor [Ho... | 770 | 0.0 |
| prf 1009174A macroglobulin alpha2 [Homo sapiens] | 765 | 0.0 |
| ref NP_036620.1 (NM_012488) alpha-2-macroglobulin [Rattus norve... | 754 | 0.0 |
| pir JC5143 alpha-macroglobulin precursor - guinea pig >gi 13040... | 736 | 0.0 |
| ref NP_075591.1 (NM_023103) alpha(1)-inhibitor 3, variant I [Ra... | 734 | 0.0 |
| sp P14046 A1I3_RAT ALPHA-1-INHIBITOR III PRECURSOR >gi 91945 pir... | 723 | 0.0 |
| ref NP_002855.1 (NM_002864) pregnancy-zone protein; Pregnancy z... | 719 | 0.0 |
| pir JC5144 murinoglobulin precursor - guinea pig | 718 | 0.0 |
| dbj BAA12317.1 (D84339) murinoglobulin [Cavia porcellus] | 715 | 0.0 |

Figure 4:

```
>ref|NP_000005.1| (NM_000014) alpha 2 macroglobulin precursor [Homo sapiens]
sp|P01023|A2MG_HUMAN Alpha-2-macroglobulin precursor (Alpha-2-M)
pir||MAHU alpha-2-macroglobulin precursor - human
gb|AAA51551.1| (M11313) alpha-2-macroglobulin precursor [Homo sapiens]
Length = 1474
```

```
Score = 770 bits (1989), Expect = 0.0
Identities = 417/917 (45%), Positives = 580/917 (62%), Gaps = 40/917 (4%)
```

```
Query: 1      QVSLGFSPSQQLPGAEVELQLQAAPGSLCALRAVDESVLLRPDRELNSRSVYGMFPFWY 60
          +V L FSPSQ LP +    L++ AAP S+CALRAVD+SVLL++PD ELS  SVY + P
Sbjct: 567    KVDLSFSPSQSLPASHAHLRVTAAPQSVCALRAVDQSVLLMKPDAELSASSVYNLLP--- 623
```

```
Query: 61     GHYPYQVAEYDQCPVSGPWDFPQPLIDPMPQGHSSQRSIIWRP-SFSEGTDLFSFFRDVG 119
          E D      GP +  Q   D + + +    I + P S +  D++SF  D+G
Sbjct: 624    -----EKDLTGFPGPLN-DQDDEDCINRHNVYINGITYTPVSSTNEKDMYSFLEDMG 674
```

```
Query: 120    LKILSNAKIKKPVDCH-----RSPE-----YSTAMGAGGGHPEAFESSTPLHQAEDSQ 168
          LK  +N+KI+KP  C          PE      + +  G  GH          P
Sbjct: 675    LKAFTNSKIRKPCKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEEP----HTET 730
```

```
Query: 169    VRQYFPETWLWDLFPIGNNSGKEAVHVTVPDAITEWKAMSFC TSQSRGFGLSPTVGLTAFK 228
          VR+YFPETW+WDL  + ++G  V  VTVPD ITEWKA +FC S+  G  G+S T  L AF+
Sbjct: 731    VRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISSTASLRAFQ 790
```

```
Query: 229    PFFVDLTLPPYSVVRGESFRLTATIFNYLKDCIRVQTDLAKSHEYQLESWADSQTSSCLCA 288
          PFFV+LT+PYSV+RGE+F L AT+ NYL  CIRV  L  S +      Q  C+CA
Sbjct: 791    PFFVELTMPYSVIRGEAFTLKATVNYLPKCIRVSVQLEASPAFLAVPVEKEQAPHCICA 850
```

```
Query: 289    DDAKTHHWNITAVKLGHINFTISTKILD SNEPCGGQKGFPQKGRSDTLIKPVLVKPEGV 348
          + +T  W +T  LG++NFT+S + L+S E CG +  VP+ GR DT+IKP+LV+PEG+
Sbjct: 851    NGRQTVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSVPEHGRKDTVIKPLLVEPEGL 910
```

```
Query: 349    LVEKTHSSLLCPKGKVASESVSLELPVDIVPDSTKAYVTVLGDI MGTALQNL DGLVQMP S 408
          E T +SLCP G  SE +SL+LP ++V +S +A V+VLDI+G+A+QN  L+QMP
Sbjct: 911    EKETTFNSLLCPGGEVSEELSLKLPNVVEESARASVSVLGDILGSAMQNTQNL LQMPY 970
```

```
Query: 409    GCGEQNMVLFAPIIYVLOYLEKAGLLTEEIRSRAVGFL EIGYQKELMYKHSNGSYSAFGE 468
          GCGEQNMVLFAP IIYVL YL +  LT E++S+A+G+L  GYQ++L YKH +GSYS FGE
Sbjct: 971    GCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGYLN TGYQRQLNYKHYDGSYTFGE 1030
```

7/24

Query: 469 RDG--NGNTWLTAFTKCFGQAQKFIFIDPKNIQDALKWMAGNQLPSGCYANVGNLLHTA 526
R G GNTWLTAFTKCFGQAQKFIFIDPKNIQDALKWMAGNQLPSGCYANVGNLLHTA 526
Sbjct: 1031 RYGRNQGNTWLTAFTVLTFAQARAYIFIDEAHITOALIWLQRQKDNGCFRSSGSLLNNA 1090

Query: 527 MKGGVDEVSLTAYVTAALLEMGKDVPDDPMVSQGLRCLKNSATST-----TNLYTQALL 580
+KGGV+DEV+L+AY+T ALLE+ V P+V L CL+++ + +++YT+ALL
Sbjct: 1091 IKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFCLESAWKTAQEGDHGSHVYTKALL 1150

Query: 581 AYIFSLAGEMDIRNILLKQLDQQAIISGESIYWSQKPTPSSNASPWSEPA--VDVELTA 638
AY F+LAG D R +LK L+++A+ S++W + P + + EP A +VE+T+
Sbjct: 1151 AYAFALAGNQDKRKEVLKSLNNEAVKKDNSVHWERPQKPKAPVGHFYEPQAPSAEVEMTS 1210

Query: 639 YALLAQLT-KPSLTQKEIAKATSIVAWLAKQHNAYGGFSSTQDTVVALQALAKY-ATTAY 696
Y LLA LT +P+ T +++ AT+IV W+ KQ NA GGFSSTQDTVVAL AL+KY A T
Sbjct: 1211 YVLLAYLTAQPAPTSEDLTSATNIVKWITKQQNAQGGFSSTQDTVVALHALSKYGAATFT 1270

Query: 697 MPSEEINLVVKSTENFQRTFNIQSVNRLVFQQDTLPNVPGMYTLEASGQGCVYVQTVLRY 756
+ + ++S+ F F + + NRL+ QQ +LP +PG Y+++ +G+GCVY+QT L+Y
Sbjct: 1271 RTGKAAQVTIQSSGTFSKFQVDNNNRLLLQQVSLPELPGEYSMKVTGEGCVYLQTSKY 1330

Query: 757 NILPPTNMKTFSLSVIEIGKARCEQPTSPRSLLTIIHTSYVGSRSSNMAIVEVKMLSGFS 816
NILP F+L V+ C++P + S +++ SY GSRS+SNMAIV+VKM+SGF
Sbjct: 1331 NILPEKEEFPPFALGVQTLQPTCDEPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFI 1390

Query: 817 PMEGTNXXXXXXPLVKKVEFGTDTLNIYLDLIKNTQTYTFTISQSVLVTNLKPATIKVY 876
P++ T V + E ++ + IYLD++ T + FT+ Q V V +LKPA +KVY
Sbjct: 1391 PLKPTVKMLERSNHVSRTEVSSNHVLIYLDKVSNQTLSLFTVLQDV PVRDLKPAIVKVY 1450

Query: 877 DYLYPDEQATIQYSDPC 893
DYY DE A +Y+ PC
Sbjct: 1451 DYYETDEFAIAEYNAPC 1467

Figure 5:

```

1 gttcctcagg coagatctga cccactggca tttattacat tttctgctaa aggagccact
      v p q a r s d p l a f i t f s a k g a t

61 ctcaacctgg aagagaggag atctgtggca atcagatcca gagagaatgt ggtcttcgta
      l n l e e r r s v a i r s r e n v v f v

121 cagactgata aacccaccta caagcctgga cagaaagtcc atatattaac attattttta
      q t d k p t y k p g q k v h i l t l f l

181 tttttatbtc agtacccagt gatcacccctt caggatccctc aaaacaatcg gattttca
      f l f q y p v i t l q d p q n n r i f q

241 aggcaaaatg tgacttctt ccgaaatatt acccaactct cgttccaaact gatttcagaa
      r q n v t s f r n i t q l s f q l i s e

301 ccaatgttg gagattactg gattgttg aaaaagaaact caagggagac agtgacacac
      p m f g d y w i v v k r n s r e t v t h

361 caatttgctg taaaagata tggctgccc aagtttgaag ttacagtcaa tgcaccacaa
      q f a v k r y v l p k f e v t v n a p q

421 acagtaacta tttcagatga tgaattccaa gtggatgtat gtgctaagta caactttggc
      t v t i s d d e f q v d v c a k y n f g

481 caacctgtgc aaggggaaac ccaaattccgg gtgtcagag agtattttc ttcaagcaat
      q p v q g e t q i r v c r e y f s s s n

541 tgtgagaaaa atgaaaatga aatatgttag caatttattg cacagttggaaatgggtgt
      c e k n e n e i c e q f i a q l e n g c

601 gtttctcaaa ttgttaatac aaaagtcttc caactctacc gttcgggatt gttcatgaca
      v s q i v n t k v f q l y r s g l f m t

661 ttcatgtcg ctgtaattgt tacagaatct gggacagttt tgcagatcag cgagaagacc
      f h v a v i v t e s g t v m q i s e k t

721 tcagtttta tcactcaatt gcttggact gtaaaactttt agaacatggta tacattctat
      s v f i t q l l g t v n f e n m d t f y

781 agaagagggta ttcttattt tgaaactctt aaattttcgatcccaataa tgtacctatg
      r r g i s y f g t l k f s d p n n v p m

841 gtgaacaagt tggcaact ggagctcaat gatgaattta taggaaattta cactacggat
      v n k l l q l e l n d e f i g n y t t d

901 gagaatggcg aagctcaatt ttccattgac acttcagaca tatttgcattcc agagttcaac
      e n g e a q f s i d t s d i f d p e f n

961 ctaaaagcca catatgtcg acctgagagc tgctatctt ccagctgggt gacgcctcag
      l k a t y v r p e s c y l p s w l t p q

1021 tacttggatg ctcacttctt agtctcacgc ttttactccc gaaccaacag cttcctgaa
      y l d a h f l v s r f y s r t n s f l k

1081 attgttccag aaccaaagca gcttgaatgt aatcaacaga agttgttac tggcattac
      i v p e p k q l e c n q q k v v t v h y

```

1141 tccctaaaca gtgaagcata tgaggatgat tccaatgtaa agttcttcta tttgatgatg
 s l n s e a y e d d s n v k f f y l m m
 1201 gtaaaaggag ctatcttact cagtgacaa aaggaaatca gaaacaaagc ctggaatgga
 v k g a i l l s g q k e i r n k a w n g
 1261 aacttctcg tcccaatcag catcagtct gatctggctc ctgcagccgt cctgttgtc
 n f s f p i s i s a d l a p a a v l f v
 1321 tatacccttc accccagtgg ggaaattgtg gctgacagtg tcagattcca gttgacaag
 y t l h p s g e i v a d s v r f q v d k
 1381 tgctttaaac acaaggtaa cataaagttc tctaacgagc agggcttacc tggttccaaat
 c f k h k v n i k f s n e q g l p g s n
 1441 gctagtctct gtcttcaagc ggcgcctgtc ttattctgtg ccctcaggcc tggatagg
 a s l c l q a a p v l f c a l r a v d r
 1501 aatgtccttc tactgaaatc tgaacaacag ctgtcagctg aaagtgtgt aacatggtt
 n v l l l k s e q q l s a e s v y n m v
 1561 ccaagtatag agccgtatgg ttatttctac catggcctca atcttgcata tggcaaggaa
 p s i e p y g y f y h g l n l d d g k e
 1621 gacccttgca ttcctcagag ggatatgtt tacaatggtt tatattacac acctgtaaac
 d p c i p q r d m f y n g l y y t p v s
 1681 aactatgggg atggagatat ctataatatt gtcaggaaca tgggtctaaa agtctttacc
 n y g d g d i y n i v r n m g l k v f t
 1741 aatctccatt accgaaaacc agaagtatgt gtgatggaga gaaggctgccc actccctaag
 n l h y r k p e v c v m e r r l p l p k
 1801 ccgcttatac tggaaacaga aaattatggt ccaatgcgtt gtgtccgtc tagaattgca
 p l y l e t e n y g p m r s v p s r i a
 1861 tgttagagggg agaatgctga ctatgtgaa caggctataa ttcaaacagt aagaacaaac
 c r g e n a d y v e q a i i q t v r t n
 1921 ttcccagaga catggatgtg ggacctcgatc agtgcgtt cctcaggctc tgccatctt
 f p e t w m w d l v s v d s s g s a n l
 1981 tcgttctca ttcctgatac gataacccaa tgggaggcaa gtggctttt tggatgggt
 s f l i p d t i t q w e a s g f c v n g
 2041 gacgttggat ttggcatttc ctctacaacc actctagaag tctcccaacc tttctttatt
 d v g f g i s s t t t l e v s q p f f i
 2101 gagattgcct cacccttttc ggttgcataa aatgaacaat ttgatttgat tgtcaatgtc
 e i a s p f s v v q n e q f d l i v n v
 2161 ttcagctacc ggaatacatg tgttagagatt tctgttcaag tggaggagtc tcagaattat
 f s y r n t c v e i s v q v e e s q n y
 2221 gaagcaaata ttcataccctt gaaaatcaat ggcagtggagg ttattcaagc tggagggagg
 e a n i h t l k i n g s e v i q a g g r
 2281 aaaacaaacg tctggactat tatacctaag aaattgggt aagtgaatat cactgttagtt
 k t n v w t i i p k k l g k v n i t v v

2341 gctgagtcca aacaaagcag tgcttgccca aatgaaggaa tggagcagca aaagctaaac
 a e s k q s s a c p n e g m e q q k l n
 2401 tggaaagaca ctgtggtcca aagcttctta gtagagcctg aaggtattga aaaggaaagg
 w k d t v v q s f l v e p e g i e k e r
 2461 acccagagtt tccttatctg tacagaaggt gccaaagcct ccaaggcaggg agttttggac
 t q s f l i c t e g a k a s k q g v l d
 2521 ttgccaaacg atgttagtaga agggtcagcc agaggcttt tcactgttgc gggggatatt
 l p n d v v e g s a r g f f t v v g d i
 2581 ctaggacttg cttgcagaa tctggttgtt ctccaaatgc cctatggaag tggagagcag
 l g l a l q n l v v l q m p y g s g e q
 2641 aatgctgccc tactagcatc tgatacttat gttctggact atctgaaatc tactgagcaa
 n a a l l a s d t y v l d y l k s t e q
 2701 ctgacagagg aagttcaatc taaggcttc ttctcttat ctaatggta tcaaaggcaa
 l t e e v q s k a f f l l s n g y q r q
 2761 ttatcttca aaaactctga tggtcctat agtgtgttt ggcagcagag tcagaaagga
 l s f k n s d g s y s v f w q q s q k g
 2821 agcatatgtg ctcttacttt taagacattg gagagaatga aaaaatatgt attcattgat
 s i c a l t f k t l e r m k k y v f i d
 2881 gaaaatgttc aaaaacagac cttaatctgg ctttcaagcc aacagaaaaac aagcggctgc
 e n v q k q t l i w l s s q q k t s g c
 2941 tttaagaatg atggccagct ttcaaccac gcctgggagg gtggagatga agaggacatt
 f k n d g q l f n h a w e g g d e e d i
 3001 tcactcactg cgtatgttgc tggatgttc ttgaagctg ggctcaattt cactttct
 s l t a y v v g m f f e a g l n f t f p
 3061 gctctacgaa acgcactctt ttgccttgaa gcggcattgg acagtgggtc cactaatggc
 a l r n a l f c l e a a l d s g v t n g
 3121 tataatcatg caattctagc ttatgctttt gccttagctg gaaaagagaa gcaagtggaa
 y n h a i l a y a f a l a g k e k q v e
 3181 tctttactcc aaaccctgga tcaatctgcc ccaaaactaa ataatgtcat ctactggaa
 s l l q t l d q s a p k l n n v i y w e
 3241 agagaaaagga aacccaagac agaagaattt ccatccctta ttccctggc accttctgtc
 r e r k p k t e e f p s f i p w a p s a
 3301 cagactgaga agagttgcta cgtgctgttgc gctgtcattt cccggaaaat tcctgaccc
 q t e k s c y v l l a v i s r k i p d l
 3361 acctatgcta gtaagattgt gcagtggctt gccaaacgga tgaattccca tggaggctt
 t y a s k i v q w l a q r m n s h g g f
 3421 tcttccaacc agggaaactgc agttgtctt ctggccataa cccgctacat aaccagggg
 s s n q e t a v c l l a i t r y i t q g
 3481 ctcttctcta aggatcaaaa cactgtcacc tttagcagtg aaggatccag tgagatttc
 l f s k d q n t v t f s s e g s s e i f

xxx = INSP087 predicted receptor binding domain.

12/24

Figure 6**a)**

| | |
|-----------------------|---|
| INSP087rbd 2243344 | ATGGTGGTTATAGATGTAATAATGCTATCAGGATTACTCCAACCATGTCATCCATTGAA ATGGTGGTTATAGATGTAATAATGCTATCAGGATTACTCCAACCATGTCATCCATTGAA ***** |
| INSP087rbd 2243344 | GAGCTTGAAAACAAGGGCCAAGTGTGAAAGACTGAAGTCAGAAATGACCATGTTCTTTG GAGCTTGAAAACAAGGGCCAAGTGTGAAAGACTGAAGTCAGAAATGACCATGTTCTTTG ***** |
| INSP087rbd 2243344 | TACTTGAAAATGGTTTGGTCGAGCAGACAGTTCCCTTTCTGTTGAGCAGAGCAAC TACTTGAAAATGGTTTGGTCGAGCAGACAGTTCCCTTTCTGTTGAGCAGAGCAAC ***** |
| INSP087rbd 2243344 | CTTGTGTTAACATTAGCCAGCCCCAGCCATGGTCTACGATTATTACGAAAAAGAAGAA CTTGTGTTAACATTAGCCAGCCCCAGCCATGGTCTACGATTACTATGAAAAAGAAGAA ***** |
| INSP087rbd 2243344 | TATGCCCTAGCTTTTACAACATCGACAGTAGTCAGTTCCAG TATGCCCTAGCTTTTACAACATCGACAGTAGTCAGTTCCAG ***** |

b)

| | |
|-----------------------|---|
| INSP087rbd 2243344 | MVVIDVKMLSGFTPMSIEELENKGQVMKTEVKNDHVLFYLENGFGRADSFPPSVEQSN MVVIDVKMLSGFTPMSIEELENKGQVMKTEVKNDHVLFYLENGFGRADSFPPSVEQSN ***** |
| INSP087rbd 2243344 | LVFNIQPAPAMVYDYYEKEEYALAFYNIDSSSVSQ LVFNIQPAPAMVYDYYEKEEYALAFYNIDSSSVSE ***** |

Figure 7:

```

1 caggtttccc ttggcttctc cccctccag cagcttccag gagcagaagt ggagctgcag
    q v s l g f s p s q q l p g a e v e l q

61 ctgcaggcag ctcccgatc cctgtgtgcg ctccggcgg tggatgagag tgtcttactg
    l q a a p g s l c a l r a v d e s v l l

121 cttaggccag acagagagct gagcaaccgc tctgtctatg ggatgtttcc attctggat
    l r p d r e l s n r s v y g m f p f w y

181 ggtcactacc cctatcaagt ggctgagat gatcagtgtc cagtgctgg cccatgggac
    g h y p y q v a e y d q c p v s g p w d

241 tttcctcagc ccctcattga cccaatgccc caagggcatt cgagccagcg ttccattatc
    f p q p l i d p m p q g h s s q r s i i

301 tggaggccct cgttctctga aggacggac ctttcagct tttccggga cgtggccctg
    w r p s f s e. g t d l f s f f r d v g l

361 aaaatactgt ccaatgccaa aatcaagaag ccagtagatt gcagtcacag atctccagaa
    k i l s n a k i k k p v d c s h r s p e

421 tacagcaactg ctatgggtgc aggccgtggc catccagagg ctttgagtc atcaactcct
    y s t a m g a g g g h p e a f e s s t p

481 ttacatcaag cagaggattc tcaggtccgc cagtagttcc cagagacctg gctctggat
    l h q a e d s q v r q y f p e t w l w d

541 ctgtttccta ttggtaactc gggaaaggag gcggtccacg tcacagttcc tgacgccatc
    l f p i g n s g k e a v h v t v p d a i

601 accgagtgga aggcgatgag tttctgcact tcccagtcaa gaggcttcgg gctttcaccc
    t e w k a m s f c t s q s r g f g l s p

661 actgttggac taactgcttt caagccgttc tttgttgcacc tgactctccc ttactcgtat
    t v g l t a f k p f f v d l t l p y s v

721 gtccgtgggg aatcctttcg tcttactgccc accatcttca attacctaaa ggattgcata
    v r g e s f r l t a t i f n y l k d c i

781 agggttcaga ctgacctggc taaatcgcat gagtaccagc tagaatcatg ggcagattct
    r v q t d l a k s h e y q l e s w a d s

841 cagacctcca gttgtctcg tgctgatgac gcaaaaaccc accactggaa catcacagct
    q t s s c l c a d d a k t h h w n i t a

901 gtcaaattgg gtcacattaa cttaacttatt agtacaaaga ttctggacag caatgaacca
    v k l g h i n f t i s t k i l d s n e p

961 tgtgggggcc agaagggtt tgttcccaa aaggccgaa gtgacacgct catcaagcc
    c g g q k g f v p q k g r s d t l i k p

1021 gttctcgta aacctgaggg agtcctggc gagaagacac acagcttatt gctgtgcacca
    v l v k p e g v l v e k t h s s l l c p

1081 aaaggaaagg tggcatctga atctgtctcc ctggagctcc cagtgacat tggccatc
    k g k v a s e s v s l e l p v d i v p d

```

1141 tcgacccaagg cttatgttac ggttctggga gacattatgg gcacagccct gcagaacctg
 s t k a y v t v l g d i m g t a l q n l
 1201 gatggctctgg tgcagatgcc cagtggctgt ggcgagcaga acatggtctt gtttgcctcc
 d g l v q m p s g c g e q n m v l f a p
 1261 atcatctatg tcttgcagta cctggagaag gcagggctgc tgacggagga gatcaggtct
 i i y v l q y l e k a g l l t e e i r s
 1321 cgggcagtgg gtttcctgga aatagggtac cagaaggagc tcatgtacaa acacagcaat
 r a v g f l e i g y q k e l m y k h s n
 1381 ggctcataca gtgccttgg ggagcgagat ggaaatggaa acacatggct gacagcgitt
 g s y s a f g e r d g n g n t w l t a f
 1441 gtcacaaaat gctttggcca agtcagaaa ttcatcttca ttatcccacaa gaacatccag
 v t k c f g q a q k f i f i d p k n i q
 1501 gatgctctca agtggatggc aggaaaccag ctccccagtg gctgctatgc caacgtggaa
 d a l k w m a g n q l p s g c y a n v g
 1561 aatctccttc acacagctat gaagggttgt gttgatgatg aggtctcctt gactgcgtat
 n l l h t a m k g g v d d e v s l t a y
 1621 gtcacagctg cattgctgga gatggaaag gatgttagatg accaatggc gagtcagggt
 v t a a l l e m g k d v d d p m v s q g
 1681 ctacggtgtc tcaagaattc ggccacctcc acgaccaacc tctacacacaca ggccctgttg
 l r c l k n s a t s t t n l y t q a l l
 1741 gcttacattt tctccctggc tggggaaatg gacatcagaa acatttcct taaacagtta
 a y i f s l a g e m d i r n i l l k q l
 1801 gatcaacagg ctatcatctc aggagaatcc atttactgga gccagaaaacc tactccatca
 d q q a i i s g e s i y w s q k p t p s
 1861 tcgaacgcca gcccttggc tgagcctgcg gctgttagatg tgaactcac agcatatgca
 s n a s p w s e p a a v d v e l t a y a
 1921 ttgttggccc agcttaccaa gcccagcctg actcaaaagg agatagcgaa ggccactagc
 l l a q l t k p s l t q k e i a k a t s
 1981 atagttggctt gggtggccaa gcaacacaat gcatatggg gcttctcttc tactcaggat
 i v a w l a k q h n a y g g f s s t q d
 2041 actgttagttg ctctccaagc tcttgccttca tatgccacta ccgcctacat gccatctgag
 t v v a l q a l a k y a t t a y m p s e
 2101 gagatcaacc tggttgtaaa atccactgag aatttccagc gcacattcaa catacagtca
 e i n l v v k s t e n f q r t f n i q s
 2161 gttaacagat tggtatttca gcaggatacc ctgcccattg tccctggaaat gtacacgttg
 v n r l v f q q d t l p n v p g m y t l
 2221 gaggcctcag gccagggctg tgtctatgtc cagacggtgt tgagatacaa tatttcctt
 e a s g q g c v y v q t v l r y n i l p
 2281 cccacaaaata tgaagacctt tagtcttagt gtggaaatag gaaaagctag atgtgagcaa
 p t n m k t f s l s v e i g k a r c e q

2341 ccgacttcac ctcgatcctt gactctcact attcacacca gttatgtggg gagccgtac
p t s p r s l t l t i h t s y v g s r s
2401 tcttcc█ata█tgcattgt ggaagtgaag atgcattatctt ggttcagtcc catggagggs
s s n m a i v e v k m l s g f s p m e g
2461 accaatcagt taattctccat gcaacccctg gtgaagaagg ttaatattgg aactgacaca
t n q l 1 1 q q p l v k k v e f g t d t
2521 cttAACattt aacttggatga gctcattaag aacactcaqa cttacaccc taccatcago
l n i y 1 d e 1 i k n t q t y t f t i s
2581 caaaactatgc tggtcaccaa tcttbaaaacca gaaaccatca aggtctatga ctactacca
q s v l v t n 1 k p a t i k v y d y y l
2641 ccagatqaac aaggcaadaat tcaqatattctt gatcccttgt aatgaggtaa gtccagcgg
p d e q a t i q y s d p c e
2701 gaaatgggtg gagttatggg tttaggtggc agaagttaaag aggagcctct tttcgagtta
2761 ctgtcattgt cttttttga gatagagtct cgcgggttg cccaggctgg agggcagtgg
2821 cggaggtt

█ = INSP088 predicted receptor binding domain.

Figure 8:**a)**

| | |
|-----------------------|---|
| INSP088rbd 4753534 | AATATGGCTATTGTGGAAGTGAAGATGCTATCTGGGTTAGTCCATGGAGGGCACCAAT AATATGGCTATTGTGGAAGTGAAGATGCTATCTGGGTTAGTCCATGGAGGGCACCAAT ***** |
| INSP088rbd 4753534 | CAGTTACTCTCCAGCAACCCCTGGTGAAGAAGGTTGAATTGGAACTGACACACTTAAC CAGTTACTCTCCAGCAACCCCTGGTGAAGAAGGTTGAATTGGAACTGACACACTTAAC ***** |
| INSP088rbd 4753534 | ATTTACTTGGATGAGCTCATTAAGAACACTCAGACTTACACCTCACCATCAGCCAAAGT ATTTACTTGGATGAGCTCATTAAGAACACTCAGACTTACACCTCACCATCAGCCAAAGT ***** |
| INSP088rbd 4753534 | GTGCTGGTCACCAACTTGAAACACCAGCAACCATCAAGGTCTATGACTACTACCTACCAAGAT GTGCTGGTCACCAACTTGAAACACCAGCAACCATCAAGGTCTATGACTACTACCTACCAAGAT ***** |
| INSP088rbd 4753534 | GAACAGGGCAACAATTCAGTATTCTGATCCCTGTGAA GAACAGGGCAACAATTCAGTATTCTGATCCCTGTGAA ***** |

b)

| | |
|-----------------------|---|
| INSP088rbd 4753534 | NMAIVEVKMLSGFSPMEGTNQLLQQPLVKKVEFGTDTLNIYLDELIKNTQTYTFTISQS NMAIVEVKMLSGFSPMEGTNQLLQQPLVKKVEFGTDTLNIYLDELIKNTQTYTFTISQS ***** |
| INSP088rbd 4753534 | VLVTNLKPATIKVYDYYLPDEQATIQYSDPCE VLVTNLKPATIKVYDYYLPDEQATIQYSDPCE ***** |

17/24

Figure 9:

Molecule: pENTR-INSPO88-sigptd-6HIS, 2615 bps DNA Circular
 File Name: pENTR-INSPO88-sigptd-6HIS-V1b.cm5

| Type | Start | End | Name | Description |
|--------|-------|------|-------------------------|----------------|
| REGION | 27 | 129 | attL1 | |
| GENE | 136 | 201 | IL12p40 signal sequence | |
| GENE | 202 | 483 | INSPO88-6HIS cds | |
| REGION | 487 | 588 | attL2 | |
| MARKER | 634 | C | | pENTR R primer |
| GENE | 710 | 1519 | KmR | |
| REGION | 1636 | 2309 | pUC ori | |
| MARKER | 2612 | | | pENTR F primer |

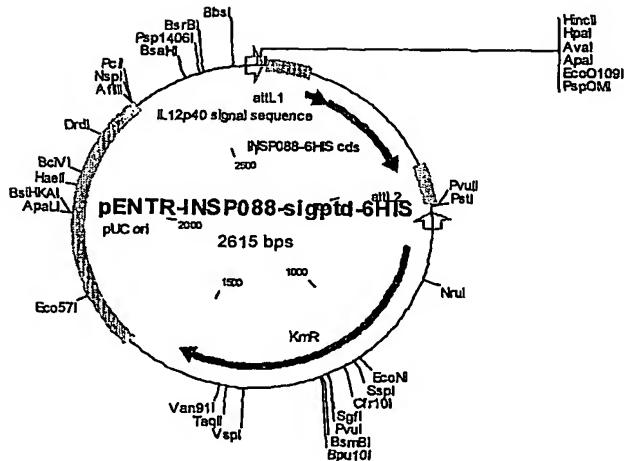


Figure 10:

Molecule: pEAK12d-INSP088-sigptd-6HIS, 7297 bps DNA Circular
File Name: pEAK12d-INSP088-sigptd-6HIS-V1b.cm5

| Type | Start | End | Name | Description |
|--------|-------|------|--------------|----------------------------|
| REGION | 2 | 595 | | pmb-ori |
| GENE | 596 | 1519 | AmpR | Ampicillin resistance gene |
| REGION | 1690 | 2795 | EF-1alpha | promoter |
| MARKER | 2703 | | | pEAK12 F primer |
| REGION | 2796 | 2845 | | MCS'' |
| REGION | 2855 | 2874 | | attB1 |
| GENE | 2888 | 2953 | IL12p40 | signal sequence |
| GENE | 2954 | 3235 | INSP088-6HIS | cds |
| REGION | 3243 | 3264 | | attB2 |
| REGION | 3270 | 3270 | | 'MCS |
| REGION | 3271 | 3699 | | poly A/splice |
| MARKER | 3385 | C | | pEAK12 R primer |
| GENE | 4318 | 3700 | C | Puromycin resistance gene |
| REGION | 4542 | 4319 | C tK | tK promoter |
| REGION | 5037 | 4543 | C Ori P | |
| GENE | 7089 | 5037 | C EBNA-1 | |
| REGION | 7090 | 7289 | sv40 | |

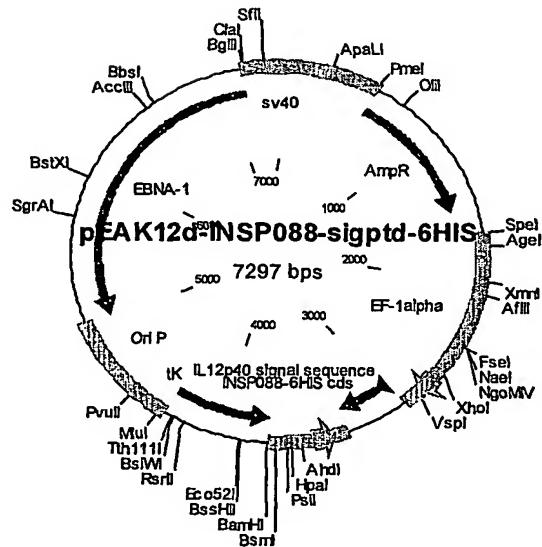


Figure 11:

1 caggttccc ttggcttctc cccctccag cagcttcag gaggcagaat ggagctgcag
 q v s l g f s p s q q l p g a e v e l q
 61 ctgcaggcag ctccggatc cctgtgtgcg ctccggcgg tggatgagag tgtcttactg
 l q a a p g s l c a l r a v d e s v l l
 121 cttaggccag acagagagct gagcaaccgc tctgtctatg ggatgttcc attctggat
 l r p d r e l s n r s v y g m f p f w y
 181 ggtcactacc cctatcaagt ggctgagtt gatcagtgtc cagtgtctgg cccatggac
 g h y p y q v a e y d q c p v s g p w d
 241 tttcctcagc ccctcattga cccaatgccc caagggcatt cgagccagcg ttccattatc
 f p q p l i d p m p q g h s s q r s i i
 301 tggaggccct cgttctctga aggcacggac cttttcaact ttttccggga cgtggccctg
 w r p s f s e g t d l f s f f r d v g l
 361 aaaatactgt ccaatgccaa aatcaagaag ccagtagatt gcagtcacag atctccagaa
 k i l s n a k i k k p v d c s h r s p e
 421 tacagcactg ctatgggtgc aggccgtgg catccagagg cttttgagtc atcaactcct
 y s t a m g a g g g h p e a f e s s t p
 481 ttacatcaag cagaggattc tcaggtccgc cagtacttcc cagagacctg gctctggat
 l h q a e d s q v r q y f p e t w l w d
 541 ctgtttccta ttggtaactc gggaaaggag gcggtccacg tcacagttcc tgacgccatc
 l f p i g n s g k e a v h v t v p d a i
 601 accgagtgga aggcgatgag tttctgcact tcccagtcaa gaggcttcgg gctttcaccc
 t e w k a m s f c t s q s r g f g l s p
 661 actgttggac taactgcttt caagccgttc tttgttggacc tgactctccc ttactcagta
 t v g l t a f k p f f v d l t l p y s v
 721 gtccgtgggg aatcccttcg tcttactgccc accatcttca attacctaaa ggattgcata
 v r g e s f r l t a t i f n y l k d c i
 781 agggttcaga ctgacctgac taaatcgcat gagtaccagc tagaatcatg ggcagattct
 r v q t d l a k s h e y q l e s w a d s
 841 cagacctcca gttgtctctg tgctgtgac gcaaaaaccc accactggaa catcacagct
 q t s s c l c a d d a k t h h w n i t a
 901 gtcaaattgg gtcacattaa cttaacttatt agtacaaaga ttctggacag caatgaacca
 v k l g h i n f t i s t k i l d s n e p
 961 tgtggggcc aqaaagggtt tggtcccaa aaggccgaa gtgacacgct catcaagcca
 c g g q k g f v p q k g r s d t l i k p
 1021 gttctcgta aacctgaggg agtcctggtg gagaagacac acagctcatt gctgtgccc
 v l v k p e g v l v e k t h s s l l c p
 1081 aaagggaaagg tggcatctga atctgtctcc ctggagctcc cagtggacat tgttcctgac
 k g k v a s e s v s l e l p v d i v p d

1141 tcgaccaagg cttatgttac gggtctggga gacattatgg qcacagccct gcagaacctg
 s t k a y v t v' l g d i m g t a l q n l
 1201 gatggtctgg tgcagatgcc cagtggctgt qgcqagcaga acatggtctt gtttgcctcc
 d g l v q m p s g c g e q n m v l f a p
 1261 atcatctatq tcttgcagta cctggagaag qcagggctgc tgacggagga gatcaggct
 i i y v l q y l e k a g l l t e e i r s
 1321 cgggcagtgg gtttccctgga aatagggtac cagaaggagc tgatgtacaa acacagcaat
 r a v g f l e i g y q k e l m y k h s n
 1381 ggctcataca gtgcctttgg ggagcqagat ggaaatggaa acacatggct gacagcggtt
 g s y s a f g e r d g n g n t w l t a f
 1441 gtcacaaaat gcttggcca agctcagaaa ttcatcttca ttgatccaa gaacatccag
 v t k c f g q a q k f i f i d p k n i q
 1501 gatgctctca agtggatggc aggaaaccag ctccccagtq gctgctatgc caacgtggg
 d a l k w m a g n q l p s g c y a n v g
 1561 aatctccttc acacagctat qaagggtggt gttgatgatg aggtctcctt qactqcgat
 n l l h t a m k g g v d d e v s l t a y
 1621 gtcacagctq cattgctggga gatgggaaag gatgtagatg accaatggt gagtcagggt
 v t a a l l e m g k d v d d p m v s q g
 1681 ctacgggtgc tcaagaattc ggccacctcc acgaccaacc tctacacaca ggccctgttq
 l r c l k n s a t s t t n l y t q a l l
 1741 gcttacattt tctccctggc tggggaaatg gacatcagaa acattctcct taaacagtt
 a y i f s l a g e m d i r n i l l k q l
 1801 gatcaacagg ctatcatctc aggagaatcc atttactgga gccagaaacc tactccatca
 d q q a i i s g e s i y w s q k p t p s
 1861 tcgaacqcca gcccttggtc tgacqctgq gctgtagatg tggactcac agcatatgca
 s n a s p w s e p a a v d v e l t a y a
 1921 ttgttggccc agcttaccaa gcccagctg actcaaaagg agatagcgaa ggccactagc
 l l a q l t k p s l t q k e i a k a t s
 1981 atagtggctt gggtggccaa gcaacacaaat gcatatgggg gcttctttc tactcaggat
 i v a w l a k q h n a y g g f s s t q d
 2041 actgttagttq ctctccaaqgc tcttgcctaa tatgccacta ccgcctacat gccatctgag
 t v v a l q a l a k y a t t a y m p s e
 2101 gagatcaacc tgggtgtaaa atccactgag aatttccagc qcacattcaa catacagtca
 e i n l v v k s t e n f q r t f n i q s
 2161 gttaacagat tggattttca gcaggatacc ctgcccattq tccctggat gtacacgttq
 v n r l v f q q d t l p n v p g m y t l
 2221 gaggcctcag gccaggctg tgtctatgtg cagacgggtgt tgagatacaa tattctccct
 e a s g q g c v y v q t v l r y n i l p
 2281 cccacaaaata tgaagacctt tagtcttagt gtggaaatag gaaaagctag atgtgagcaa
 p t n m k t f s l s v e i g k a r c e q

2341 ccgacttcac ctcgatcctt gactctact attcacacca gttatgtggg gagccgtac
 p t s p r s l t l t i h t s y v g s r s

2401 tcttccaaaa tggctattgt ggaactqaag atgctatctg lggttcaagtcc catggaggcd
 INSP088-CP1
 s s n m a i v e v k m l s g f s p m e g

2461 accaaatcaat tacttctca gcaacccctg iqtgaagaagg ttgaatttgg aactgacaca
 t n q l l l q q p l v k k v e f g t d t

2521 cttaaacatcc acttggatqa gctcattaag iaacactcaga cttacacaccc caccatgagc
 l n i y l d e l i k n t q t y t f t i s

2581 caaagtgtgc tggcaccaa cttgaaacca gcaaccatca laggcttatqa ctactaccta
 q s v l v t n l k p a t i k v y d y y l

2641 ccagatgaaat aggcaacaat tcagtattct gateccctgtg aatgaggtaa gtccagcgg
 INSP088-CP2
 p d e q a t i q y s d p c e

2701 gaaatgggtg gagttatggg ttagggtggc agaagttaag aggagcctct tttcgagtta
 2761 ctgtcattgt ctttttttga gatagagtct cgcgggttg cccaggctgg agggcagtgg
 2821 cgagggtt

xxx = INSP088 predicted receptor binding domain.

xxx = INSP088 predicted macroglobulin domain.

→ Position and sense of PCR primers-

Figure 12:

1 gtggtcatcc agaggcttt gagtcataa ctccttaca tcaagcagag gattctcagg
 INSP088-CP3 →
 g h p e a f e s s t p l h q a e d s q

61 tccgcagta cttcccagag acctggctct gggatctgtt tcctattgggt aactcgggga
 v r q y f p e t w l w d l f p i g n s g

121 aggaggcggt ccacgtcaca gttcctgacg ccatcaccga gtggaaggcg atgagttct
 k e a v h v t v p d a i t e w k a m s f

181 gcacttccca gtcaagaggc ttcgggcttt caccactgt tggactaact gctttcaagc
 c t s q s r g f g l s p t v g l t a f k

241 cattctttgt tgacctgact ctcccttact cagtagtccg tggggaatcc tttcgctta
 p f f v d l t l p y s v v r g e s f r l

301 ctgccaccat cttcaattac ctaaaggatt gcatcagggt tcagactgac ctggctaaat
 t a t i f n y l k d c i r v q t d l a k

361 cgcatgagta ccagctagaa tcatggcag attctcagac ctccagttgt ctctgtgctg
 s h e y q l e s w a d s q t s s c l c a

421 atgaagcaaa aacccaccac tggAACatca cagctgtcaa atgggtcac attaacttta
 d e a k t h h w n i t a v k l g h i n f

481 ctatttagtac aaagatttg gacagcaatg aaccatgtgg gggccagaag gggttgttc
 t i s t k i l d s n e p c g g q k g f v

541 cccaaaaggc ccgaagtgac acgctcatca agccagttct cgtcaaacct gagggagtcc
 p q k g r s d t l i k p v l v k p e g v

601 tggTggagaa gacacacagc tcattgtgt gcccAAAagg aaaggTggca tctgaatctg
 l v e k t h s s l l c p k g k v a s e s

661 tctccctgga gctcccagtg gacattgttc ctgactcgac caaggcttat gttacggttc
 v s l e l p v d i v p d s t k a y v t v

721 tgggagacat tatgggcaca gcccgcaga acctggatgg tctggcag atgcccagtg
 l g d i m g t a l q n l d g l v q m p s

781 gctgtggcga gcagaacatg gtctgtttg ctcccatcat ctatgtctt cagtaactgg
 g c g e q n m v l f a p i i y v l q y l

841 agaaggcagg gctgctgacg gaggagatca ggtctcgcc agtgggtttc ctggaaatag
 e k a g l l t e e i r s r a v g f l e i

901 ggtaccagaa ggagctgatg tacaaacaca gcaatggctc atacagtgcc tttggggagc
 g y q k e l m y k h s n g s y s a f g e

961 gagatggaaa tggaaacaca tggctgacag cgtttgcac aaaatgctt ggccaaagctc
 r d g n g n t w l t a f v t k c f g q a

1021 agaaattcat cttcattgtat cccaaagaaca tccaggatgc tctcaagtgg atggcaggaa
 q k f i f i d p k n i q d a l k w m a g

1081 accagctccc cagtggctgc tatgccaacg tggaaatct ccttcacaca gctatgaagg
 n q l p s g c y a n v g n l l h t a m k

1141 gtgggtttga tgatgaggc tccttgactg cgtatgtcac agctgcattg ctggagatgg
 g g v d d e v s l t a y v t a a l l e m
 1201 gaaaggatgt agatgaccca atggtagtc agggctatg gtgtctcaag aattcggcca
 g k d v d d p m v s q g l w c l k n s a
 1261 cctccacgac caacctctac acacaggccc tggtggctta cattttctcc ctggctgggg
 t s t t n l y t q a l l a y i f s l a g
 1321 aaatggacat cagaaacatt ctccctaaac agttagatca acaggctatc atctcaggag
 e m d i r n i l l k q l d q q a i i s g
 1381 aatccattta ctggagccag aaacctactc catcatcgaa cgccagccct tggcttgagc
 e s i y w s q k p t p s s n a s p w s e
 1441 ctgcggctgt agatgtggaa ctcacagcat atgcattgtt ggcccagctt accaagccca
 p a a v d v e l t a y a l l a q l t k p
 1501 gcctgactca aaaggagata gcgaaggcca ctagcatagt ggcttggttg gccaagcaac
 s l t q k e i a k a t s i v a w l a k q
 1561 gcaatgcata tggggcttc tcttctactc aggatactgt agttgctctc caagctttg
 r n a y g g f s s t q d t v v a l q a l
 1621 ccaaataatgc cactaccgcc tacgtgccat ctgaggagat caacctgggtt gtaaaatcca
 a k y a t t a y v p s e e i n l v v k s
 1681 ctgagaattt ccagcgcaca ttcaacatac agtcagttaa cagattggta tttcagcagg
 t e n f q r t f n i q s v n r l v f q q
 1741 ataccctgcc caatgtccct ggaatgtaca cgttggaggc ctcaggccag ggctgtgtct
 d t l p n v p g m y t l e a s g q g c v
 1801 atgtgcagac ggtgttgaga tacaatattc tccctccac aaatatgaag acctttatgc
 y v q t v l r y n i l p p t n m k t f s
 1861 tttagtgtgaa aataggaaaa gctagatgtg agcaaccgac ttcacctcga tccttgactc
 l s v e i g k a r c e q p t s p r s l t
 1921 tcactattca caccagttat gtggggagcc gtagctcttc caatatggct attgtggaaag
 l t i h t s y v g s r s s s n m a i v e
 1981 tgaagatgtc atctgggttc agtcccatgg agggcaccaa tcagttactt ctccagcaac
 v k m l s g f s p m e g t n q l l l q q
 2041 ccctggtaa gaagggtgaa tttggactg acacacttaa catttacttg gatgagctca
 p l v k k v e f g t d t l n i y l d e l
 2101 ttaagaacac tcagacttac accttcacca tcagccaaag tggctggtc accaacttga
 i k n t q t y t f t i s q s v l v t n l
 2161 aaccagcaac catcaaggc tatgactact acctaccaga tgaacaggca acaattcagt
 k p a t i k v y d y y l p d e q a t i q
 2221 attctgatcc ctgtgaatga ggtaagtgcc agc

INSP088-CP2

Position and sense of PCR primers →

Figure 13:

Molecule: pCR4-TOPO-INSPO88-CP2/-CP3, 6209 bps DNA Circular

| Type | Start | End | Name | Description |
|--------|-------|------|---------|----------------------------------|
| MARKER | 205 | | M13R | M13 rev priming site |
| MARKER | 243 | | T3 | T3 priming site |
| REGION | 295 | 2546 | Insert | |
| MARKER | 295 | | CP3 | INSP088-CP3 cloning primer |
| GENE | 336 | 2531 | cds | INSP088 macroglobulin domain cds |
| MARKER | 794 | | SP1 | INSP088-SP1 sequencing primer |
| MARKER | 1221 | | SP2 | INSP088-SP2 sequencing primer |
| MARKER | 1689 | | SP3 | INSP088-SP3 sequencing primer |
| MARKER | 2240 | | CP1 | INSP088-CP1 amplification primer |
| MARKER | 2527 | C | CP2 | INSP088-CP2 cloning primer |
| MARKER | 2532 | | stop | INSP088 stop codon |
| MARKER | 2580 | C | T7 | T7 priming site |
| MARKER | 2607 | C | M13F | M13 for priming site |
| GENE | 3411 | 4205 | Kan | Kanamycin resistance gene ORF |
| GENE | 4409 | 5269 | Amp | Ampicillin resistance gene ORF |
| REGION | 5414 | 6087 | pUC ori | pUC origin |

